8

Information Item

Educational Policy and Programs Committee

Progress Report on the Commission's 2001 College Eligibility Study

In what many consider its most important work, the California Postsecondary Education Commission (and a predecessor agency) has, over the past several decades, provided periodic estimates of the proportion of public high school graduates eligible to attend California's public universities in comparison to the admissions guidelines in the California Master Plan for Higher Education.

This information has proved critical in State educational program and facilities planning. Past Commission eligibility studies were published by the Commission in 1976, 1985, 1988, 1992 and 1997. Commission staff has completed the initial year of work on the 2001 College Eligibility Study and reports on those activities in this information item. Included is information about the broad-based Eligibility Study Task Force and the work completed on the study's sampling design specifications.

The remaining steps necessary to complete the study for submission as information item in December 2002 are also set forth.

Presenters: Stacy Wilson.

Progress Report on the Commission's 2001 College Eligibility Study

Introduction

Since 1960, the California Legislature has requested the California Postsecondary Education, and its immediate predecessor, the Coordinating Council for Higher Education, to conduct eligibility studies that examine changes in student academic preparation for college. A principal goal of previous eligibility studies has been to obtain an unbiased estimate, with high statistical precision, of the proportion of public high school graduates eligible to attend the California State University (CSU) and the University of California (UC) in comparison to freshman admission guidelines established in the California Master Plan for Higher Education.

Display 1 on page 2 provides a summary of statewide eligibility results for the period 1955 to 1996, which includes results from the founding study undertaken by the *Committee for the Restudy of the Needs of California in Higher Education*. Clear understanding of those results have been necessary for programmatic planning purposes and for assessing the extent to which construction of new campus facilities and off-campus centers will be needed to guarantee college opportunities for future generations of college-bound seniors.

During the past year, Commission staff has been at work on a new eligibility study that will estimate and assess the eligibility status of the 2001 class of public high school graduates. This progress report describes some of the key pre-planning activities, outlined in Appendix A, completed during the initial year of the current study.

It is anticipated that work on this report will continue through most of 2002, with a draft report to be submitted at the Commission's December 2002, meeting. It would follow that the report would be presented for adoption by the Commission in February 2003.

The State eligibility study task force and sampling design committee

College eligibility studies are among the most complex empirical research efforts undertaken by the Commission on behalf of the State of California. The success of past studies has rested greatly on the active participation of the University of California, the California State University, and the State Department of Education through the appointment and committee work of a Statewide Eligibility Task Force. The current task force also includes representatives from the Department of Finance, the Association of Independent California Colleges and Universities, and the academic subcommittees that oversee admission requirements at the State's public universities.

DISPLAY 1 *Estimated Freshman Eligibility Rates for the California State University and the University of California, 1955 to 1996*

Master Plan Guidelines Established in 1960	**Year**	**CSU** 33.3%	**UC** 12.5%
Responsible Agency			
Committee for the Restudy of the Needs of California in Higher Education	1955	44.0	15.0
Master Plan Survey Team	1961	43.4	14.8
Coordinating Council for Higher Education	1966	35.2	14.6
California Postsecondary Education Commission	1975	35.0	14.8
California Postsecondary Education Commission	1983	29.6	13.2
California Postsecondary Education Commission	1986	27.5	14.1
California Postsecondary Education Commission	1990	34.6	12.3
California Postsecondary Education Commission	1996	29.6	11.1

Source: California Postsecondary Education Commission.

A critical task force undertaking was the appointment of a Sampling Design Technical Subcommittee to develop valid and reliable design specifications for sampling the transcript records of the 2001 class of public high school graduates. Given the complexity of policy issues that have emerged since the 1996 study, including the University of California's new path to eligibility, called *Eligibility in the Local Context*, a limited amount of consulting funds (approximately $30,000) were made available to hire a team of statisticians to assist the technical subcommittee.

Following a national search, consulting services were awarded to three individuals: (1) a former UC Berkeley mathematical statistician, who has also served as a principal research scientist for the Educational Testing Service, (2) a principal survey researcher form the Bureau of Labor Statistics, located in Washington D.C., and (3) a social scientist/statistician from California State University, Sacramento. The Eligibility Task Force adopted unanimously the 2001 statewide sampling design proposed by the technical subcommittee. That design is discussed in detail later in this progress report.

Selected issues and goals addressed by the task force

Among the study goals adopted by the task force was a commitment to again obtain eligibility estimates for three demographic subgroups at minimum: (1) gender, (2) ethnic-racial group, and (3) geographic region. As noted in the *2001 Eligibility Study Prospectus*, the 1983 study was the first to derive eligibility estimates for these subgroups.

In order to ensure that the results of the current study will be directly comparable to previous results, the task force took appropriate steps to maintain comparable sampling precision. This will require that the over-

all statewide eligibility rates for the State University and the University of California be estimated with a sampling error no greater then plus/minus one percentage point, and that eligibility rates for all subpopulations be estimated with sampling errors no greater than plus/minus three percentage points.

A key issue with which the present and previous task forces on eligibility have struggled is the need to obtain a more representative sampling of African American graduates from both high- and low-performing schools. Addressing the issue of African-American representation has been difficult because that particular ethnic-racial group is not as evenly distributed across the state as are the other racial groups.

Thus, any proposed design that calls for applying a single sampling rate across all schools would almost certainly result in a less-than optional sample of African-American graduates from high-performing schools, where their representation tends to be low. Accordingly, the task force directed its technical design subcommittee to use the statewide estimated 1996 UC Eligibility Medium as a stratification variable to better allow for a desirable sampling of African-American graduates from a cross-section of public high schools with respect to college eligibility performance.

Another issue of central concern to the Task Force involved the level of support that previous studies had received from the Educational Testing Service (ETS) and the American College Testing Service (ACT). Those two testing services administer the Scholastic Assessment Tests and ACT Examination, respectively. In the past, both *ETS* and *ACT* matched its test file against the Commission's eligibility file and supplied test information for matched records. This process has been deemed vital for ensuring that the study will have all necessary test information, even though, in most cases, that information is reported on the student's final transcript by the local high school.

Therefore, the task force initiated discussions with both testing services to clarify and enhance the manner in which eligibility records are to be matched against *EST* and *ACT* data files during the new eligibility study.

Overview of the 2001 sampling design

Display 2 highlights the design features that will be used to sample the transcript records of the 2001 class of public high school graduates. The sampling frame consists of all public comprehensive, continuation, alternative, and adult school that reported any graduates during year 2001. As mentioned earlier, a primary goal of the design is to provide an unbiased estimate with high statistical accuracy of the statewide proportion of graduates eligibility to attend the University of California and the California State University. This goal is achieved through a random probability sampling strategy that provides for a sufficient sample size within each California public high school.

The design was also developed with an interest to obtain accurate eligibility estimates by gender, ethnic-racial group, and region. In particular, special effort was directed towards ensuring a more representative sampling of African-American graduates from both high-performing and low-performing public high schools. (White, Asian, and Latino graduates are distributed, for the most part, more evenly across both high- and low-performing schools, so it is quite certain that representative samples of those ethnic groups will be obtained without adopting a specific strategy beyond a simple random selection approach within each school.)

A 10% sampling rate will be used for all public high schools with an estimated 1996 UC eligibility rate above the statewide median and whose African-American graduates represent between 3% and 7% of the school's 2001graduating class. A census of African-American transcripts will be requested from a school if its African-American representation is less than 3% of the its graduating class. This should help ensure an adequate sampling of African-American graduates from public schools that have had historically high college eligibility rates.

DISPLAY 2 Specific Sampling Deign Features of the 2001 College Eligibility Study

<u>*Sampling Frame*</u>. All public comprehensive, continuation, alternative, and adult schools that had any graduates during the 2000-01 academic year will be sampled.
<u>*Eligibility Estimates*</u>. As directed by the Eligibility Study Task Force, the 2001 study will obtain UC and CSU eligibility estimates for the overall statewide mean and for three subgroups: (1) gender, (2) ethnic-racial group, and (3) geographic region.
<u>*Stratification Features*</u>. The *1996 UC Estimated Eligibility Median* (.0769%) will be used as the primary stratification variable. Thus, graduates will be grouped by two levels: (1) those graduating from a school that had a UC eligibility rate above the median and (2) and those graduating from a school that had a 1996 UC eligibility rate below the statewide median. Within each median level, graduates are grouped naturally by school attended, so *public high school*, in effect, also represents a non-overlapping stratum.
<u>*Sampling Rates*</u>. A 10% sampling rate will be used for all public high schools that had an estimated 1996 UC Eligibility Rate above the statewide median and whose African-American graduates represent between 3% and 7% of the school's 2000 graduating class. A census of African-American transcripts will be requested from schools above the UC Eligibility Median if this ethnic group represents 3% or less of the graduating class. A 5% sampling rate be applied to all remaining schools. The proposed sampling rates translate to an overall sampling rate of approximately 6%, which is essentially the statewide rate adopted in the 1996 study.

Display 3 shows the estimated sample size, by design source, for graduates of comprehensive public high schools that results from the adopted design. The 15,577 transcripts intended to be sampled translates to an overall sampling rate of approximately 6%, which is essentially the same as the rate adopted in the 1996 study. Statistical weighting procedures, as well as appropriate procedures for estimating sampling precision, will be used to ensure unbiased eligibility estimates. Appropriate weighting procedures are necessary whenever differential sampling rates are used to estimate a population parameter of interest.

In order to establish the precision of the eligibility estimates it is first necessary to calculate the variance in college eligibility summed across schools. The square root of that mathematical expression provides a measure of precision for the overall UC and CSU statewide eligibility estimates. Two slightly different methods will be used to establish the level of sampling precision for the subgroups. A technical discussion of those methods are provided in Appendix B.

DISPLAY 3 Estimated Sample Size for Comprehensive Public High Schools

African-Am. students included with certainty	All students from 10% sampling rate	All students from 5% sampling rate			**TOTAL**
		Above UC Eligibility Median & 0%-3% African-Am. students	Above UC Median & GT 7% African-Am. students	Below UC Eligibility Median	
1,051	3,009	3,765	1,859	5,893	**15,577**

Design items yet to be resolved

In order to estimate the entire pool of graduates eligible to attend the University of California, the overall statewide estimate will need to be statistically adjusted to account for the University's new eligibility path, called *Eligibility in the Local Context*. In general, this path explicitly recognizes that student academic achievement is tied, in numerous ways, to the level of academic support resources available to students across socioeconomic school districts of California. Accordingly, the top 4% of college-bound seniors of each high school who complete all required course requirements will be considered UC eligible.

More specifically, students must have completed the following 11 academic units by the end of their junior year: (a) one unit of history/social science, (b) three units of English, (c) three units of mathematics, (d) one unit of laboratory science, (e) one unit of foreign language, and (f) and two units chosen from among the other subject requirements. Students' high school rankings will be based on their grade point averages in those courses. At the present, no agreement has been reached on the most appropriate method for incorporating the *ELC* path into the statewide de-

sign. It is anticipated that the issue will be resolved by the March 2002 of the Eligibility Study Task Force meeting.

The 2001 study will be the first of the present century undertaken in an era of expanded public accountability in schooling. Perhaps not since the early 1980s has school reform and public accountability been so prevalent at the national, state, and local levels in an effort to address student academic achievement and college preparedness. The task force discussed various options for relating college eligibility data to the *State's Public Schools Accountability ACT of 1999*. Although consensus has yet to be reached, a promising approach that has been received favorably is to incorporate the Department of Education's School Characteristics Index (SCI) as a post-hoc design feature of the eligibility study.

The SCI is composed of indicators related to *socioeconomic status, English-language learning status, ethnic-racial group, teacher credentialing, school mobility, and class size*. The index reflects an explicit acknowledgement that there are social factors beyond a school's control that influence individual academic achievement.

Because schools in different socioeconomic and demographic surroundings often face different challenges and opportunities in meeting the needs of learners they serve, California law requires the Department of Education to compare a school's academic performance in relation to that of other schools similarly situated. Such comparisons enable school districts, in developing their local school improvement plans, to identify promising practices by looking to higher performing schools with similar characteristics.

Because it was found that SCI explains more than 80% of the variation in statewide high school Academic Performance Index (API) scores, it is almost certain that SCI would also explain a significant proportion of variation in college eligibility. This means that a number of salient policy issues could potentially be addressed by using SCI as a post-hoc design feature of the current eligibility study. A specific recommendation will be considered by March 2002.

Display 4 provides an updated timeline of remaining issues and activities to be addressed or initiated during year two of the eligibility study.

DISPLAY 4 Proposed 2nd Year Timeline for the 2001 College Eligibility Study Of California Public High School Graduates

Proposed Timeline **Year 2 (2002)**	**Activity**
January	Mail sampling instructions to all public comprehensive, continuation, alternative, and adult schools.
	Select a state vendor to scan transcripts and create electronic files.
	Receive and process transcripts, and initiate first follow-up request for transcripts from non-responding schools.
Feb. thru March	Resolve design issues related to *UC's Eligibility in the Local Context* admission path and the Department of Education's *School Characteristic Index* (SCI).
	Develop key-entry screens, invoice-processing procedures, and procedures for responding to inquiries from high school staff; generate unique identifier codes.
April thru May	Process in-coming transcripts, verify sample number, assign code numbers, cover-up all personally identifiable data, and key-enter basic demographic data.
	Photocopy transcripts and submit to state vendor for scanning and creating electronic files. Send electronic files to the CSU and UC.
	Implement second full-scale non-response follow-up. Match test information against data received from ETS and ACT.
	Process eligibility data submitted by the CSU and UC. Match eligibility data to the CPEC demographic file and identify any missing transcripts. Resolve mismatches in coding between UC and the CSU.
June thru September	Calculate eligibility rates by gender, ethnic-racial group, and geographic region. Verify rates with UC and the CSU. Resolve any discrepancies.
	Analyze eligibility data and write final report.
December	Submit Eligibility Report as an Information Agenda Item.

Appendix A

Year-One Pre-planning Activities Completed related to the Commission's 2001

College Eligibility Study

2001 Eligibility Study Prospectus. A detailed prospectus was developed by staff that described the purpose, background, tasks timelines, and general research design issues of the current study involving the 2001 class of public high school graduates. A briefing on the prospectus was heard by the Commission at its February 2001 meeting.

Committee Appointments. A Statewide Eligibility Task Force was appointed to provide comprehensive, broad-based advice on implementing the 2001 study. The Task Force includes representatives from higher education, K-12 education, the Demographic research Unit of the State Department of Finance, and representatives from the UC and CSU academic senate committees that oversee freshman admission requirements. A Sampling Design Subcommittee was appointed to resolve technical design issues and to develop specific sampling specifications.

Statistical Consultant Search. A national search was undertaken to identify a team of statistical/survey consultants to assist the Commission's Sampling Design Subcommittee in developing a valid and reliable stratified methodology for sampling the transcript records of public high school graduates. Consulting contracts were awarded to three individuals: (1) a former UC Berkeley mathematical statistician, (2) a principal survey statistician from the Bureau of Labor Statistics, and a Social Scientists/Statistician from California State University, Sacramento.

Public High School Notification Letters. A letter announcing the 2001 eligibility study was sent to every public comprehensive high school, continuation, alternative, and adult schools that had any graduates during the 1999-00 school year.

Memorandum of Understanding (MOU). A MOU between the California Postsecondary Education Commission and the University of California and the California State University was agreed upon and signed, defining respective roles, tasks, procedures, and responsibilities for conducting the 2001 eligibility study.

Adoption of the 2001 Study Design. The technical subcommittee developed complete sampling design specifications for the 2001 College Eligibility Study. The design was adopted unanimously by the Task Force at its December 2001 meeting.

Appendix B

Notes on estimating eligibility rates and variances for ethnic and gender subgroups in the 2001 College Eligibility Study

Juliet P. Shaffer

(Notes prepared for meeting of Dec. 18, 2001)

Since the design specifies sampling rates for every school, it is a stratified design with schools as strata. The standard method of estimating the eligibility rate and the variance for stratified samples is appropriate for the whole graduating population.

Standard estimator:

Let N_i be the total number of graduates in School i, n_i be the sampled number of graduates, and p_i be the sample proportion of eligible graduates. Providing $n_i > 1$, $[p_i(1-p_i)/(n_i - 1)](1 - n_i/N_i)$ is an unbiased estimate of the school variance. Adding the $N_i p_i$ values over schools and dividing by N, where $N = \sum_i(N_i)$, gives an unbiased estimate of the eligibility over all schools, and

$$\sum_i[(N_i/N)^2[p_i(1-p_i)/(n_i-1)](1-n_i/N_i)]$$

is an estimate of the variance. There will be a slight downward bias in the variance estimate, since schools with a sample size of 1 will not yield a variance estimate. These will be very small schools, so the bias should be minimal.

Let A represent an ethnic or gender subgroup. Let N_{Ai} be the total number of graduates in School i who are in subgroup A, n_{Ai} be the sampled number of graduates in School i who are in subgroup A, and p_{Ai} be the sample proportion of eligible graduates in school i who are in subgroup A. In many schools, the sample may include either 0 or 1 student in the sample who is in subgroup A. If it is 0, there will not be an estimate for that subgroup in that school. Thus, if the same estimation method is used for subgroup A that is used for the total group, the overall estimate will be biased towards the eligibility rates in large schools with a large proportion of students in subgroup A. Furthermore, in estimating variance, schools with either 0 or 1 students in the sample in subgroup A will have no variance estimates. Since the estimates would be added over all schools with eligibility estimates, the bias in the estimates of both total eligibility and its variance could be large for a subgroup.

Schools with large minority proportions are likely to be both larger and less affluent than other schools, and the students in those schools in all subgroups are likely to have lower eligibility rates. Thus, the estimates of eligibility for minority groups are likely to be biased downward. Because school distribution is most uneven for African-American students, this bias is most likely to affect their eligibility estimate. This is the reason the recommended survey design includes certainty sampling for African-Americans in schools with estimated 1996 eligibility rate above the median in which they comprise 3 % or less of the graduating class, and a relatively high overall rate (10 %) for schools with estimated 1996 eligibility rate above the median in which African- Americans comprise 3 to 7 % of graduates.

In addition to the increased sampling rate, a different method of estimation for subgroups should give a less-biased estimate of subgroup eligibility, as described below.

Alternative estimator:
There are three sampling rates for African-Americans (100 % sampling, 10 % sampling, and 5 % sampling) and two rates for other subgroups (10 % and 5 %). Consider a single subgroup A, and students in A only. Let S_j represent the set of schools with sampling rate j. Let N_{Aj} be the total number of subgroup-A students from all schools in S_j, n_{Aj} be the total number of those students in the sample in those schools, and let e_{Aj} be the number of students in the sample from those schools who are eligible. Then $N_{Aj}p_{Aj}$, where $p_{Aj} = e_{Aj}/n_{Aj}$, is an estimate of the total number of eligible students in the schools with sampling rate j. Add the estimated totals $N_{Aj}p_{Aj}$ over the different sampling-rate groups to get an estimate of the total number of eligible students in subgroup A, and divide by the total number of graduating students in subgroup A, N_A, to get an estimate of the eligibility rate for that subgroup. This estimator should be less biased for subgroups than the standard estimator.

An upwardly biased estimator of the variance of the estimated rate based on this alternative estimator is

$$\sum_j (N_{Aj}^2/N_A^2)[p_{Aj}(1-p_{Aj})/(n_{Aj}-1)](1-n_{Aj}/N_{Aj}),$$

(variance conditional on observed sample sizes).

This is upwardly biased because p_{Aj} is an estimate of a weighted average eligibility rate across the schools with sampling-rate j, and the variance of an average proportion is greater than the average of the variances.

A downwardly biased estimator of the variance of this total can be derived as follows.

Let e_{Aij}, n_{Aij} and N_{Aij} be the total number of eligible students in the sample, the total sample size, and the total population size, respectively, in school i in S_j, and let $p_{Aij} = e_{Aij}/n_{Aij}$ be the proportion of eligible students in school i. Then the estimated eligibility rate p_{Aj} can be written as

$$p_{Aj} = \sum_i (n_{Aij}/n_{Aj})p_{Aij},$$

and the associated estimated total over all sampling groups is

$$\sum_j N_{Aj}p_{Aj}.$$

The (conditional) variance of the estimated rate is

$$(N_A)^{-2} \sum_j N_{Aj}^2 \sum_i [(n_{Aij}/n_{Aj})^2 p_{Aij}(1-p_{Aij})/(n_{Aij}-1)](1-n_{Aij}/N_{Aij}).$$

Since there will be no estimates for the schools i with n_i equal to 0 or 1, this estimate will be downwardly biased.

If the two variance estimators are not very far apart, the upper one will be suitable. Otherwise, intermediate estimates are possible.